TITLE:

Small molecule binding to a G-hairpin and a G-triplex: a new insight into anticancer drug design targeting G-rich regions.

AUTHOR(S):

Rajendran, Arivazhagan; Endo, Masayuki; Hidaka, Kumi; Teulade-Fichou, Marie-Paule; Mergny, Jean-Louis; Sugiyama, Hiroshi

CITATION:

Rajendran, Arivazhagan ...[et al]. Small molecule binding to a G-hairpin and a G-triplex: a new insight into anticancer drug design targeting G-rich regions.. Chemical communications 2015, 51(44): 9181-9184

ISSUE DATE:

2015-04-13

URL:

http://hdl.handle.net/2433/201620

RIGHT:

This journal is © The Royal Society of Chemistry 2015.; The full-text file will be made open to the public on 13 April 2016 in accordance with publisher's 'Terms and Conditions for Self-Archiving'.; この論文は出版社版でありません。引用の際には出版社版をご確認ご利用ください。; This is not the published version. Please cite only the published version.

京都大学学術情報リポジトリ
KURENAI 紅
Kyoto University Research Information Repository

京都大学 KYOTO UNIVERSITY
A Self-archived copy in
Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp
京都大学学術情報リポジトリ KURENAI 紅 Kyoto University Research Information Repository

Journal Name

RSCPublishing

COMMUNICATION

Small molecule binding to G-hairpin and G-triplex: A new insight in anticancer drug design targeting G-rich regions†

Cite this: DOI: 10.1039/x0xx00000x

Received 00th Month 2015,
Accepted 00th Month 2015

DOI: 10.1039/x0xx00000x

www.rsc.org/

Arivazhagan Rajendran,[‡a] Masayuki Endo,*[bc] Kumi Hidaka,[a] Marie-Paule Teulade-Fichou,[d] Jean-Louis Mergny[e] and Hiroshi Sugiyama*[abc]

Although a number of G-quadruplex binders and chaperones have been reported to date, the molecular mechanism of quadruplex-drug interaction has never been investigated. For instance, the interaction between the quadruplex binders and the in-pathway intermediates such as G-hairpin and G-triplex is not reported so far. To gain new insights on quadruplex–drug interactions, we captured the solution-state structures of the complexes between a drug-like small molecule and G-hairpin/G-triplex. Our results indicated that the investigated ligand initially binds to the intermediates and induces stepwise folding into a quadruplex.

G-quadruplex-forming sequences are attractive therapeutic targets because of the high abundance of G-rich DNA regions in the human genome. Among the G-rich regions, telomeres have attracted particular interest as anticancer therapy targets because this region is responsible for the stability of the human chromosome. Telomere is a repetitive nucleotide sequence of TTAGGG that has a duplex region followed by single-stranded 3′-overhang of up to few hundred nucleobases at each end of a chromosome.[1] It is a noncoding region that protects the genetic data, making it possible for cells to divide, and it plays a vital role in both aging and cancer. It is composed of a nucleoprotein complex called shelterin, which protects the telomere ends from the DNA-damage-response machinery.[2] G-quadruplex-based anticancer therapy depends on the development and application of drugs that can replace the shelterin complex competitively by inducing the quadruplex structure and thereby lead to telomere dysfunction, which culminates in cell death.[3] Several quadruplex binders of natural origin[4–6] and synthetic molecules[7–10] or metal complexes,[11] have been recognized. Their selectivity towards the quadruplex over B-form DNA was studied under bulk conditions. However, to the best of our knowledge, the molecular mechanism of quadruplex–drug interaction has never been investigated either experimentally or by theoretical methods. Thus, the interaction between the quadruplex-binders/chaperones and the quadruplex intermediates such as G-hairpin and G-triplex is still unknown at molecular level. Indeed, even theoretical discussions on this aspect of drug binding to these in-pathway intermediates are notably absent from the literature. This is mainly because the involvement of such intermediates in the folding process of a quadruplex is still not well understood. There have been an increasing number of experimental[12–20] and theoretical[21–25] studies on these intermediates in quadruplex folding; however, these studies only assumed their involvement. We have recently demonstrated the direct and single-molecule visualization of the solution-state structures of these intermediates (induced by salt[20]) using DNA origami[26–30] and high-speed atomic force microscopy (HS-AFM).[31–36] Here, we present new insights into the binding of drug–quadruplex intermediates and propose the molecular mechanisms involved in the drug-induced quadruplex folding. By using our single-molecule method, we present the structures of drug–intermediate complexes with nanometer resolution. Other methods failed to provide concrete evidence for the presence of these intermediates and could not be used even to hypothesize such drug–intermediate interactions. The ligand that was used for our studies was the bisquinolinium pyridine dicarboxamide bearing a linker terminated by biotin (PDC-biotin, Fig. 1a).[37]

A DNA origami frame (Fig. 1b)[28–29] was used to structurally and stoichiometrically control the DNA sequences,[38] and the drug–intermediate binding for tetramolecular antiparallel and (3+1)-type structures (in which three G-tracts are present in one strand and one G-tract in another strand) were investigated.[20] For the formation of the tetramolecular G-hairpin, we adopted two unique Watson–Crick duplexes each containing six contiguous Gs opposite to a nick in the complementary strand (see, Fig. 1c and Fig. S1, ESI†). Thus, only two G-repeat strands were used out of the four participating strands, which could lead to the formation of G-hairpin structure specifically. In the case of the tetramolecular G-triplex, the top duplex contained six G-G mismatches in the middle, whereas the bottom duplex contained six contiguous Gs opposite to a nick in the complementary strand (Fig. 1d). This leaves only three strands with G-repeat sequences in the middle, which in turn can produce a G-triplex structure specifically. To bring the duplexes closer and to promote the formation of the desired intermediates, we incorporated structural flexibility into the strands by increasing their length. As a result, the lengths of the top and bottom duplexes used were 67 and 77 bp respectively, whereas the length between the two connecting sites in the origami was 64 bp.

The (3+1)-type G-hairpin was formed by adopting two duplexes, each containing a single-stranded overhang with GGGGGGTTN (where N = A or T) sequence (Fig. 1e and Fig. S2, ESI†). In the case of the (3+1)-type G-triplex, the top duplex contained two repeats of GGGGGGTTN sequence, whereas the bottom duplex contained one such sequence (Fig. 1f). We also investigated the (3+1)-type structures that contained three contiguous Gs, as it is naturally abundant (Fig. S3, ESI†). All the duplex DNAs used here contained single-stranded regions, each 16 bases in length at both termini, which are needed for their attachment inside the origami frame through complementary base pairing. Small-molecule-induced conformational switching was observed by using HS-AFM by

京都大学 KYOTO UNIVERSITY
A Self-archived copy in
Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp
京都大学学術情報リポジトリ KURENAI 紅 Kyoto University Research Information Repository

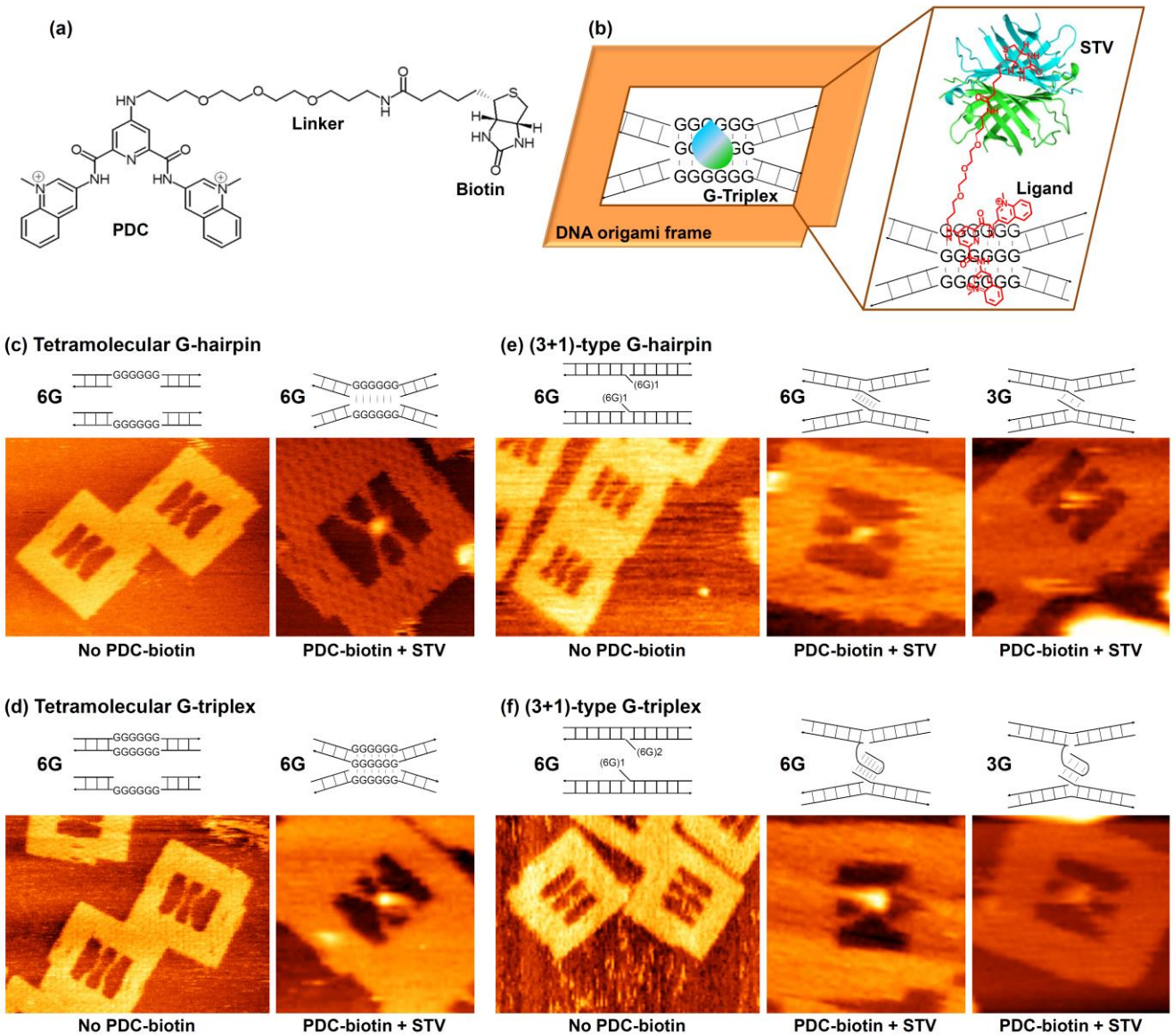


Fig. 1. (a) Line drawing of the structure of PDC-biotin ligand. (b) Schematic explanation of the DNA origami method for the analysis of ligand-induced formation of intermediate structures. The immobilization of STV to the biotin present in the ligand is also shown. Representative zoom-in AFM images of the DNA origami frame with incorporated duplexes for the tetramolecular antiparallel (c and d) and (3+1)-type (e and f) structures. Parallel and X-shape of the incorporated duplexes indicate the unstructured and characteristic intermediate structures, respectively. The bright spot at the middle of X-shape indicates the STV-biotin present in the ligand binding. For comparison, the images recorded in the absence of ligand are taken from our parallel study, reference [20]. [M13mp18] = 10 nM; [Staples] = 40 nM; [Tris-HCl] = 20 mM, pH 7.6; [$MgCl_2$] = 5 mM (no ligand) 10 mM (with ligand); [EDTA] = 1 mM; [KCl] = 0 mM; [PDC-biotin] = 1 μM; [STV] = 0.2 μM. Image size: 300 × 225 nm (no ligand) and 125 × 125 nm (with ligand).

monitoring the topological changes of the incorporated duplexes from parallel (no structure) to X-shaped (formation of a notable intermediate structure). The presence of the ligand in the intermediate structure was confirmed by the localization of streptavidin (STV) using STV-biotin (biotin present in the ligand) binding as a pixel-enhancing marker in the AFM image.

Regarding the general experimental procedure, we have prepared the DNA origami and attached the duplexes of interest in each case. The origami assembly was then purified by passing through a Sephacryl S-300 filtration column to remove the excess staples and unattached duplexes. PDC-biotin ligand was then added to the purified origami assembly, which was then immobilized on a mica surface. After gently washing the surface to remove the unbound origami assembly, STV was added on the surface and the system was incubated for 5 min. Imaging was carried out under PDC-biotin-free buffer after washing the mica surface to remove excess STV. For the detailed experimental procedure, see the ESI†.

Initially, we investigated the tetramolecular hairpin structure. The G-repeat sequences are not able to form the hairpin structure in the absence of quadruplex inducers such as PDC-biotin and salts (K^+ and Mg^{2+}), although a minor but significant amount (27%) of hairpin structure was formed even in the presence of 5 mM Mg^{2+}. Thus, a majority of the incorporated strands adopted a parallel form inside the origami frame (Fig. 1c). Interestingly, addition of PDC-biotin induced topological changes of the duplexes from parallel to X-shape by bringing the duplexes closer at the G-repeat region. Furthermore, the yield of the X-shape in this case was doubled when compared with that obtained under PDC-biotin-free condition. A representative zoom-in AFM image is given in Fig. 1c, and the zoom-out image in Fig. S4, ESI†. The calculated yields in each case

This journal is © The Royal Society of Chemistry 2015

京都大学 KYOTO UNIVERSITY
A Self-archived copy in
Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp
京都大学学術情報リポジトリ KURENAI 紅 Kyoto University Research Information Repository

Table 1. The % yield of X-shapes calculated for the G-repeats containing samples.

Sequences	G-Hairpin		G-Triplex		G-Quadruplex	
	No ligand[a]	With ligand	No ligand[a]	With ligand	No ligand	With ligand
Tetramolecular antiparallel structures						
6 G-repeat	27%	54 (15)%	24%	44 (34)%	18%[b]	69 (21)%[c]
Counted origami	416	174	153	149	331	329
(3+1)-type structures						
6 G-repeat	24%	43 (12)%	29%	55 (30)%	26%[a]	64 (33)%
Counted origami	532	112	346	159	406	137
3 G-repeat	6%	23 (10)%	11%	29 (10)%	16%[a]	40 (16)%
Counted origami	345	154	432	205	331	138

[a,b,c]For comparison, these data are taken from our parallel studies, references [20], [38] and [39] respectively. Numbers in the parenthesis indicate the yield of X-shapes with bound STV. [Tris-HCl] = 20 mM, pH 7.6; [$MgCl_2$] = 5 mM (no ligand) or 10 mM (with ligand); [EDTA] = 1 mM; [KCl] = 0 mM; [PDC-biotin] = 1 μM; [STV] = 0.2 μM.

are summarized in Table 1. This topological change indicated that the ligand induces the formation of a tetramolecular antiparallel hairpin structure. The localization of STV in the center of the X-shape, as indicated by the bright spot in the AFM image, further confirmed that the ligand is present in the G-rich core, and consequently provided evidence for the ligand-hairpin binding. Among the four participating strands, only two strands contained the G-repeats and thus the incorporated strands can only form the hairpin structure, ignoring the possibility of other types of structure such as G-triplex and G-quadruplex. A similar trend was also observed for the tetramolecular antiparallel triplex structure (Fig. 1d and Fig. S4, ESI†). In this case, 24% of strands formed X-shape in the absence of ligand, and addition of ligand induced the formation of X-shape in as many instances as 44%.

In a similar way, we have tested the (3+1)-type hairpin structures

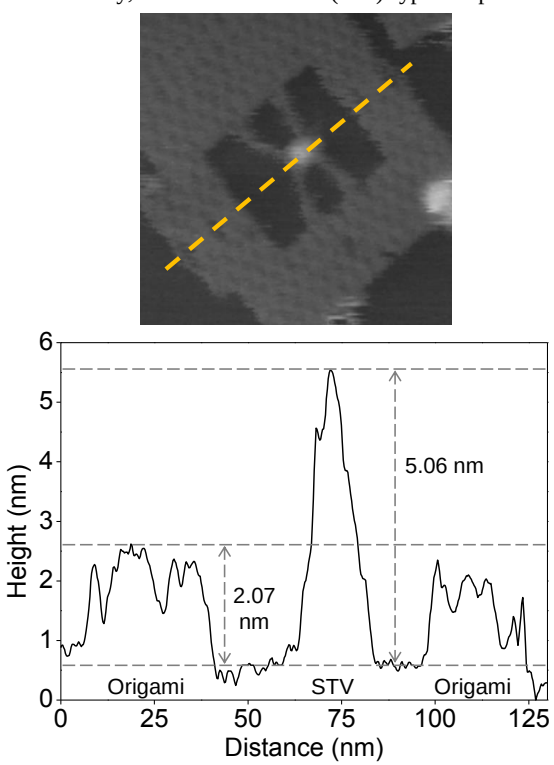


Fig. 2. The graph shows a representative height profile estimated from the image shown above. The yellow broken line in the image indicates the location where the height was measured.

(see Fig. 1e and Fig. S5 in the ESI†). For structures with six G-repeats, about 24% of the duplexes adopted parallel form in the absence of PDC-biotin, whereas this amount increased to 43% when PDC-biotin was added. Sequences containing three Gs were also tested, and the obtained X-shapes in the absence and presence of the ligand were 6 and 23%, respectively. These observations indicated that the ligand induces the (3+1)-type hairpin structure with a reasonable yield. It is worth mentioning here that three contiguous Gs avoid the potential caveat that may be possible with the relatively long six G-repeats. In the latter case, if a zero nucleotide chain reversal loop is feasible, the quadruplex may still be formed with two G3 runs from the same strand. However, it is not possible with the shorter three contiguous Gs, and thus evidencing the ligand-induced formation of G-hairpin.

The (3+1)-type triplexes were then investigated (Fig. 1f and Fig. S5, ESI†). As can be seen from Table 1, 29% of structures adopted X-shape for sequences with six G-repeats in the absence of PDC-biotin ligand and this amount was increased to 55% when the origami assembly was incubated with the ligand. Similar conformational changes were also noticed with sequences containing three G-repeats. In this case, 11 and 29% of the structures adopted X-shape in the absence and in the presence of ligand, respectively (Fig. 1f and Fig. S6, ESI†). In all cases, the localization of STV in the center of the X-shape indicated clearly that the ligand is present in the G-rich core, and consequently evidenced ligand–intermediate binding. For comparison, the yields of G-quadruplex structures in each case are listed in Table 1. In two of the three systems, the yields of the ligand-induced structures follow the order: hairpin < triplex < quadruplex, reflecting the energetics of the structures, with lowest energy conformation for the quadruplex and highest energy structure for the hairpin.

The binding of STV-biotin in the ligand was also characterized by analyzing the height profile. A representative graph of the height analysis of the protein particle inside an origami frame is given in Fig. 2. The estimated height of the origami frame and STV are 2.07 and 5.06 nm, respectively. These values are in good agreement with the theoretically expected values of 2 and 5 nm, respectively for the origami and STV.[39-40]

Sequences with the G to T mutation were also tested in selected cases to check whether the formation of X-shape was due to the formation of a characteristic structure or whether the strands simply make physical contact for a short time. The mutated sequences failed to adopt X-shape and exhibited 100% parallel form, which indicated that the X-shape formed in case of G-repeat sequences is due to the formation of a notable intermediate structure (data not shown).

京都大学 KYOTO UNIVERSITY
A Self-archived copy in
Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp
京都大学学術情報リポジトリ KURENAI 紅 Kyoto University Research Information Repository

Moreover, we have previously reported that STV alone can neither bind nor induce the formation of X-shape in G-rich sequences, which demonstrates further that formation of X-shape is due to the ligand-induced formation of the intermediate structure.[39]

In concluding remarks, by using DNA origami as a novel scaffold to control the strand polarity, stoichiometry, and number of G-repeats or G-repeat containing strands, we have investigated the formation of G-quadruplex intermediates such as G-hairpin and G-triplex structures induced by a G-quadruplex-binding PDC-biotin ligand. This ligand successfully bound to the G-repeat regions and induced the formation of intermediate structures for both tetramolecular and (3+1)-type systems. Our studies strongly support the assumptions that the G-hairpin and G-triplex are formed as in-pathway intermediates of G-quadruplex folding (irrespective of whether it is induced by salt[20] or a ligand). Furthermore, our HS-AFM investigation could capture the solution-state structures of the intermediate–ligand complex with nanometer precision, while no parallel report exists in the literature. Moreover, we would like to point out that not even a theoretical discussion exists in the literature on drug binding to these intermediates. Thus, this communication constitutes the first report of research into an unprecedented area of intermediate–ligand binding. Based on these results, we could make a general hypothesis that the G-quadruplex-binding ligands are not merely quadruplex binders, but they would rather initially bind to the in-pathway intermediates and induce stepwise folding into a quadruplex structure. These mechanistic investigations shed light on drug–intermediate binding and could help in the development of novel anticancer drugs targeting G-rich regions. Regarding the demerits of this study, it is difficult to apply these results directly to the human genome inside cells because such conditions are completely different. Furthermore, the spatial resolution of the analysis of the intermediate–ligand complex should be improved to few Angstroms.

We sincerely thank for the CREST grant from the Japan Science and Technology Corporation (JST), grants from the WPI program (WPI-iCeMS, Kyoto University) and JSPS KAKENHI (grant numbers 24310097, 24225005, 24104002 and 26620133). Financial supports from The Mitsubishi Foundation and The Sekisui Chemical Grant Program to M.E. are also acknowledged.

Notes and references

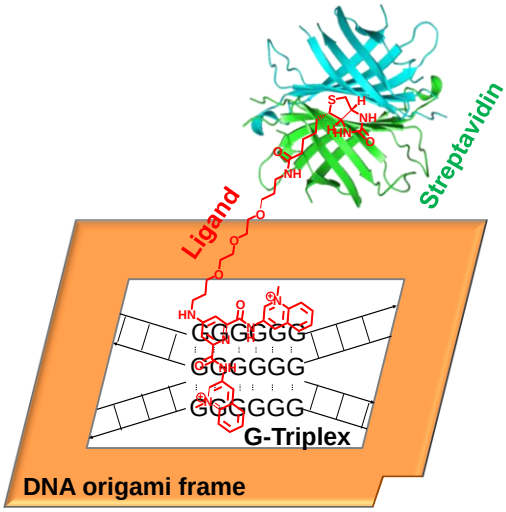
[a] *Department of Chemistry, Graduate School of Science, Kyoto University, Kitashirakawa-oiwakecho, Sakyo-ku, Kyoto 606-8502, Japan. Email: hs@kuchem.kyoto-u.ac.jp*
[b] *Institute for Integrated Cell-Material Sciences (WPI-iCeMS), Kyoto University, Yoshida-ushinomiyacho, Sakyo-ku, Kyoto 606-8501, Japan. Email: endo@kuchem.kyoto-u.ac.jp*
[c] *CREST, JST, Sanbancho, Chiyoda-ku, Tokyo 102-0075, Japan.*
[d] *Institut Curie, UMR 176 CNRS, Campus Universitaire Paris-Sud, 91405 Orsay, France.*
[e] *Univ. Bordeaux, INSERM, U869, ARNA Laboratory, 2 rue Robert Escarpit, Pessac, F-33607, France.*
[‡] *Current address: Institute of Advanced Energy, Kyoto University, Gokasho, Uji-shi, Kyoto-fu, 611-0011, Japan.*
† Electronic Supplementary Information (ESI) available. See DOI: 10.1039/c000000x/

1 V. L. Makarov, Y. Hirose, J. P. Langmore, *Cell*, 1997, **88**, 657-666.
2 T. d. Lange, *Gen. Dev.*, 2005, **19**, 2100-2110.
3 S. Muller, D. A. Sanders, M. D. Antonio, S. Matsis, J.-F. Riou, R. Rodriguez, S. Balasubramanian, *Org. Biomol. Chem.*, 2012, **10**, 6537-6546.
4 M. Franceschin, L. Rossetti, A. D'Ambrosio, S. Schirripa, A. Bianco, G. Ortaggi, M. Savino, C. Schultes, S. Neidle, *Bioorg. Med. Chem. Lett.*, 2006, 16, 1707-1711.
5 L. Guittat, P. Alberti, F. Rosu, S. V. Miert, E. Thetiot, L. Pieters, V. Gabelica, E. D. Pauw, A. Ottaviani, J. F. Riou, J.-L. Mergny, *Biochimie*, 2003, **85**, 535-547.
6 M. Y. Kim, H. Vankayalapati, S. Kazuo, K. Wierzba, L. H. Hurley, *J. Am. Chem. Soc.*, 2002, **124**, 2098-2099.
7 D. Y. Sun, B. Thompson, B. E. Cathers, M. Salazar, S. M. Kerwin, J. O. Trent, T. C. Jenkins, S. Neidle, L. H. Hurley, *J. Med. Chem.*, 1997, **40**, 2113-2116.
8 F. X. G. Han, R. T. Wheelhouse, L. H. Hurley, *J. Am. Chem. Soc.*, 1999, **121**, 3561-3570.
9 S. M. Gowan, R. Heald, M. F. G. Stevens, L. R. Kelland, *Mol. Pharmacol.*, 2001, **60**, 981-988.
10 J. T. Kern, S. M. Kerwin, *Bioorg. Med. Chem. Lett.*, 2002, **12**, 3395-3398.
11 S. N. Georgiades, N. H. A. Karim, K. Suntharalingam, R. Vilar, *Angew. Chem. Int. Ed.*, 2010, **49**, 4020-4034.
12 D. Koirala, T. Mashimo, Y. Sannohe, Z. Yu, H. Mao, H. Sugiyama, *Chem. Commun.*, 2012, **48**, 2006-2008.
13 M. Boncina, J. Lah, I. Prislan, G. Vesnaver, *J. Am. Chem. Soc.*, 2012, **134**, 9657-9663.
14 R. D. Gray, R. Buscaglia, J. B. Chaires, *J. Am. Chem. Soc.*, 2012, **134**, 16834-16844.
15 A. Y. Q. Zhang, S. Balasubramanian, *J. Am. Chem. Soc.*, 2012, **134**, 19297-19308.
16 W. Li, X.-M. Hou, P.-Y. Wang, X.-G. Xi, M. Li, *J. Am. Chem. Soc.*, 2013, **135**, 6423-6426.
17 V. Limongelli, S. D. Tito, L. Cerofolini, M. Fragai, B. Pagano, R. Trotta, S. Cosconati, L. Marinelli, E. Novellino, I. Bertini, A. Randazzo, C. Luchinat, M. Parrinello, *Angew. Chem. Int. Ed.*, 2013, **52**, 2269-2273.
18 N. An, A. M. Fleming, C. J. Burrows, *J. Am. Chem. Soc.*, 2013, **135**, 8562-8570.
19 P. M. Yangyuoru, A. Y. Q. Zhang, Z. Shi, D. Koirala, S. Balasubramanian, H. Mao, *ChemBioChem*, 2013, **14**, 1931-1935.
20 A. Rajendran, M. Endo. K. Hidaka, H. Sugiyama, *Angew. Chem. Int. Ed.*, 2014, **53**, 4107-4112.
21 R. Stefl, T. E. Cheatham III, N. Spackova, E. Fadrna, I. Berger, J. Koca, J. Sponer, *Biophys. J.*, 2003, **85**, 1787-1804.
22 T. Mashimo, H. Yagi, Y. Sannohe, A. Rajendran, H. Sugiyama, *J. Am. Chem. Soc.*, 2010, **132**, 14910-14918.
23 P. Stadlbauer, M. Krepl, T. E. Cheatham III, J. Koca, J. Sponer, *Nucleic Acids Res.*, 2013, **41**, 7128-7143.
24 V. Gabelica, *Biochimie*, 2014, **105**, 1-3.
25 P. Stadlbauer, L. Trantirek, T. E. Cheatham III, J. Koca, J. Sponer, *Biochimie*, 2014, **105**, 22-35.
26 P. W. K. Rothemund, *Nature*, 2006, **440**, 297-302.
27 A. Rajendran, M. Endo, H. Sugiyama, *Angew. Chem. Int. Ed.*, 2012, **51**, 874-890.
28 M. Endo, Y. Katsuda, K. Hidaka, H. Sugiyama, *J. Am. Chem. Soc.*, 2010, **132**, 1592-1597.
29 M. Endo, Y. Katsuda, K. Hidaka, H. Sugiyama, *Angew. Chem. Int. Ed.*, 2010, **49**, 9412-9416.
30 A. Rajendran, M. Endo, H. Sugiyama, *Curr. Protoc. Nucleic Acid Chem.*, 2012, **48**, 12.9.1-12.9.18.
31 A. Rajendran, M. Endo, Y. Katsuda, K. Hidaka, H. Sugiyama, *ACS Nano*, 2011, **5**, 665-671.
32 A. Rajendran, M. Endo, H. Sugiyama, *Chem. Rev.*, 2014, **114**, 1493-1520.
33 A. Rajendran, M. Endo, K. Hidaka, P. L. T. Tran, J.-L. Mergny, R. J. Gorelick, H. Sugiyama, *J. Am. Chem. Soc.*, 2013, **135**, 18575-18585.
34 A. Rajendran, M. Endo, K. Hidaka, H. Sugiyama, *J. Am. Chem. Soc.*, 2013, **135**, 1117-1123.
35 M. Endo, H. Sugiyama, *Acc. Chem. Res.*, 2014, **47**, 1645-1653.
36 T. Ando, T. Uchihashi, S. Scheuring, *Chem. Rev.*, 2014, **114**, 3120-3188.
37 A. Renaud de la Faverie, F. Hamon, C. D. Primo, E. Largy, E. Dausse, L. Delauriere, C. Landras-Guetta, J.-J. Toulme, M.-P. Teulade-Fichou, J.-L. Mergny, *Biochimie*, 2011, **93**, 1357-1367.
38 A. Rajendran, M. Endo, K. Hidaka, P. L. T. Tran, J.-L. Mergny, H. Sugiyama, *Nucleic Acids Res.*, 2013, **41**, 8738-8747.
39 A. Rajendran, M. Endo, K. Hidaka, P. L. T. Tran, M.-P. Teulade-Fichou, J.-L. Mergny, H. Sugiyama, *RSC Adv.*, 2014, **4**, 6346-6355.
40 K. Numajiri, M. Kimura, A. Kuzuya and M. Komiyama, *Chem. Commun.*, 2010, **46**, 5127-5129.

This journal is © The Royal Society of Chemistry 2015

京都大学 KYOTO UNIVERSITY
A Self-archived copy in
Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp
京都大学学術情報リポジトリ KURENAI 紅 Kyoto University Research Information Repository

TOC Graphics:



To gain new insights on G-quadruplex-drug interactions, we captured the solution-state structures of the complexes between a drug-like small molecule and G-hairpin/G-triplex. Our results indicated that G4-ligands initially bind to the intermediates and induce stepwise folding into a quadruplex.

京都大学 KYOTO UNIVERSITY
A Self-archived copy in
Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp
京都大学学術情報リポジトリ KURENAI 紅 Kyoto University Research Information Repository

Electronic Supplementary Information (ESI)

Small molecule binding to G-hairpin and G-triplex: A new insight in anticancer drug design targeting G-rich regions

Arivazhagan Rajendran,[‡a] Masayuki Endo,*[bc] Kumi Hidaka,[a] Marie-Paule Teulade-Fichou,[d] Jean-Louis Mergny[e] and Hiroshi Sugiyama*[abc]

[a] *Department of Chemistry, Graduate School of Science, Kyoto University, Kitashirakawa-oiwakecho, Sakyo-ku, Kyoto 606-8502, Japan.*
[b] *Institute for Integrated Cell-Material Sciences (WPI-iCeMS), Kyoto University, Yoshida-ushinomiyacho, Sakyo-ku, Kyoto 606-8501, Japan.*
[c] *CREST, JST, Sanbancho, Chiyoda-ku, Tokyo 102-0075, Japan.*
[d] *Institut Curie, UMR 176 CNRS, Campus Universitaire Paris-Sud, 91405 Orsay, France.*
[e] *Univ. Bordeaux, INSERM, U869, ARNA Laboratory, 2 rue Robert Escarpit, Pessac, F-33607, France.*
[‡] *Current address: Institute of Advanced Energy, Kyoto University, Gokasho, Uji-shi, Kyoto-fu, 611-0011, Japan.*

EXPERIMENTAL SECTION

Chemicals and reagents. Tris-HCl, EDTA, and $MgCl_2$ were purchased from Nacalai Tesque, Inc. (Kyoto, Japan). STV was purchased from Sigma-Aldrich (Japan). PDC-biotin ligand was prepared based on the synthetic procedures that we have reported recently (Faverie *et al.*, *Biochimie*, **2011**, *93*, 1357.). Single-stranded M13mp18 DNA was obtained from New England Biolabs, Inc. (Ipswich, MA). The staple strands (most of them are 32-mer) for the fabrication of the DNA origami frame, and the oligomers for the formation of intermediate structures were received from Sigma Genosys (Hokkaido, Japan). The gel-filtration column and sephacryl S-300 were purchased from Bio-Rad Laboratories, Inc. (Hercules, CA) and GE Healthcare UK Ltd. (Buckinghamshire, UK), respectively. Water was deionized (≥18.0 MΩ cm specific resistance at 25 °C) by a Milli-Q system (Millipore Corp., Bedford, MA).

Preparation of the origami frame and incorporation of the duplexes. Origami frame was prepared by annealing the solution of M13mp18 DNA (final concentration of 0.01 μM), staple DNAs (0.04 μM), Tris-HCl (20 mM, pH 7.6), EDTA (1 mM) and $MgCl_2$ (5 or 10 mM) from 85 to 15 °C at a rate of -1.0 °C/min. The duplex DNAs (final concentration of 0.1 μM each) were also prepared using the same condition with that of the origami frame. 10-fold excess of each duplex was then mixed with the origami frame. Self-assembly of these duplexes inside the origami frame was carried out by re-annealing the solution from 50 to 15 °C at a rate of -1.0 °C/min. The duplexes incorporated origami was purified using sephacryl S-300 gel-filtration column before HS-AFM imaging. For the experiments in the presence of PDC-biotin, the ligand (1 μM of final concentration) was added to the origami solution before immobilization on the mica surface. For the design of origami frame and the sequence of staple strands, see our previous publication (Endo *et al.*, *J. Am. Chem. Soc.*, **2010**, *132*, 1592.).

AFM imaging. AFM images were recorded using a fast-scanning AFM system (Nano Live Vision, RIBM Co. Ltd., Tsukuba, Japan) with a silicon nitride cantilever (resonant frequency 1.0-2.0 MHz, spring constant 0.1-0.3 N/m, EBD tip radius <15 nm, Olympus BL-AC10EGS-A2). 2 μL of sample was adsorbed onto a freshly cleaved mica plate (ϕ 1.5 mm, RIBM Co. Ltd., Tsukuba, Japan) for 5 min at room temperature and then the surface was gently washed 3-5 times using 20 mM Tris-HCl buffer solution with the same concentration of salt in which origami was prepared. For the experiments in the presence of streptavidin (STV), STV (0.2 μM) was added on the mica surface and incubated for 5 min. Images were recorded after gently washing the surface. Scanning was performed using the tapping mode. All images reported here were recorded with an image acquisition speed of 0.2 frame/s. The yield calculations of the parallel and X-shapes were carried out by counting the shapes in the AFM images. Note, the term “parallel-shape” doesn’t mean the strand polarity such as parallel G-quadruplex, it rather represents the parallel orientation of the incorporated duplexes in the AFM topographic image.

京都大学 KYOTO UNIVERSITY
A Self-archived copy in
Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp
京都大学学術情報リポジトリ KURENAI 紅 Kyoto University Research Information Repository

1) Tetramolecular G-hairpin: Six contiguous Gs

Top duplex:
5′-CTGTAGCT CATCATGT GGAGACTCTAGAGTGTTCCTGATGGCCGTA**GGGGGG**TCAAGGCGGTGGGTGCGCGTTGCTCCTCACT **GAACACCC TGAACAAA-3′**
3′-CCTCTGAGATCTCACAAGGACTACCGGCAT AGTTCCGCCACCCACGCGCAACGAGGAGTGA-5′
Bottom duplex:
5′- GAGGAAGTGAGGAGCAACGCGCACCCACCGCCTTA TCACGGCCATCAGGAACACTCTAGAGTCTCCGCTCG-3′
3′-CTAAGAGA ACAAACGA CTCCTTCACTCCTCGTTGCGCGTGGGTGGCGGAAT **GGGGGG** AGTGCCGGTAGTCCTTGTGAGATCTCAGAGGCGAGC **TACAACAA ATAACAGC-5′**

2) Tetramolecular G-triplex: Six contiguous Gs

Top duplex:
5′-CTGTAGCT CATCATGT GGAGACTCTAGAGTGTTCCTGATGGCCGTA**GGGGGG**TCAAGGCGGTGGGTGCGCGTTGCTCCTCACT **GAACACCC TGAACAAA-3′**
3′-CCTCTGAGATCTCACAAGGACTACCGGCAT**GGGGGG**AGTTCCGCCACCCACGCGCAACGAGGAGTGA-5′
Bottom duplex:
5′- GAGGAAGTGAGGAGCAACGCGCACCCACCGCCTTA TCACGGCCATCAGGAACACTCTAGAGTCTCCGCTCG-3′
3′-CTAAGAGA ACAAACGA CTCCTTCACTCCTCGTTGCGCGTGGGTGGCGGAAT **GGGGGG** AGTGCCGGTAGTCCTTGTGAGATCTCAGAGGCGAGC **TACAACAA ATAACAGC-5′**

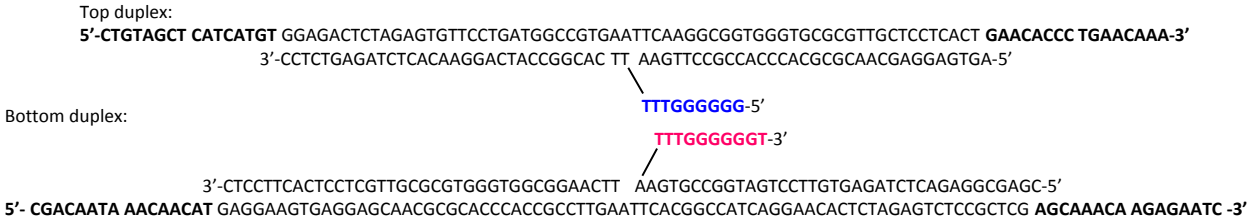
3) Tetramolecular G-quadruplex: Six contiguous Gs

Top duplex:
5′-CTGTAGCT CATCATGT GGAGACTCTAGAGTGTTCCTGATGGCCGTA**GGGGGG**TCAAGGCGGTGGGTGCGCGTTGCTCCTCACT **GAACACCC TGAACAAA-3′**
3′-CCTCTGAGATCTCACAAGGACTACCGGCAT**GGGGGG**AGTTCCGCCACCCACGCGCAACGAGGAGTGA-5′
Bottom duplex:
5′-CGACAATA AACAACAT GAGGAAGTGAGGAGCAACGCGCACCCACCGCCTTA**GGGGGG**TCACGGCCATCAGGAACACTCTAGAGTCTCCGCTCG **AGCAAACA AGAGAATC-3′**
3′-CTCCTTCACTCCTCGTTGCGCGTGGGTGGCGGAAT**GGGGGG**AGTGCCGGTAGTCCTTGTGAGATCTCAGAGGCGAGC-5′

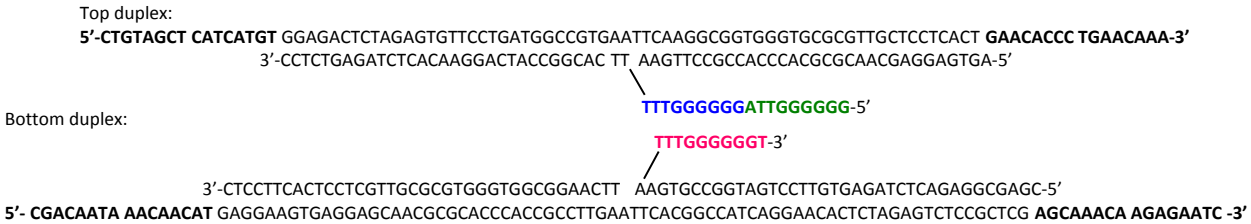
Fig. S1. The DNA sequences used for the preparation of tetramolecular antiparallel G-hairpin, G-triplex and G-quadruplexes. The colored regions at the middle are the G-repeat sequences that form the notable structure. The bold letter regions at both the termini indicate the single-stranded regions that are needed to attach these duplexes inside the origami frame. This figure is taken from the supporting information of our recent report [Rajendran *et al.*, *Angew. Chem. Int. Ed.*, **2014**, *53*, 4107.] as we have used the same sequences in this study.

京都大学 KYOTO UNIVERSITY
A Self-archived copy in
Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp
京都大学学術情報リポジトリ KURENAI 紅 Kyoto University Research Information Repository

1) (3+1)-type G-hairpin: Six contiguous Gs



2) (3+1)-type G-triplex: Six contiguous Gs



3) (3+1)-type G-quadruplex: Six contiguous Gs

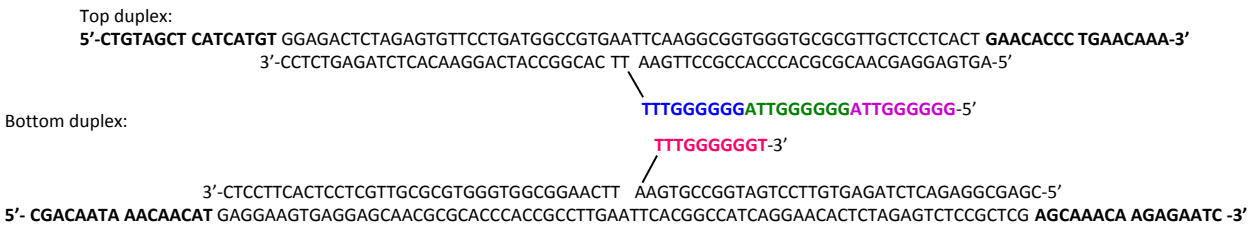
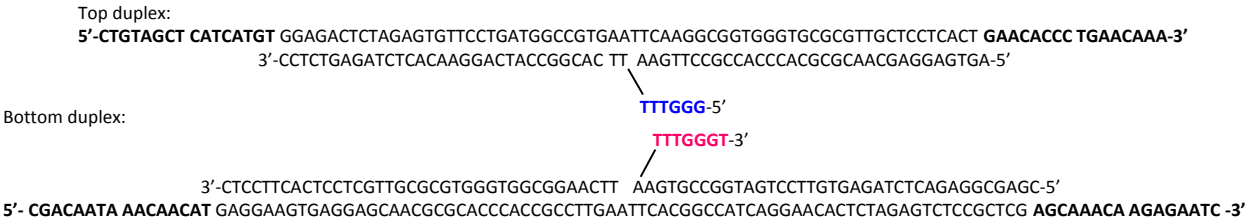


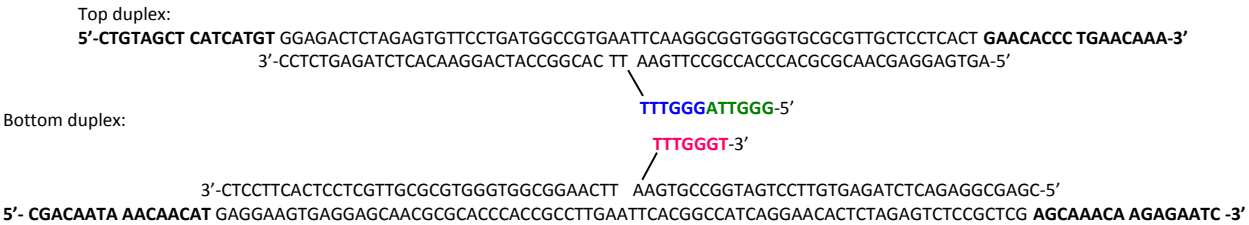
Fig. S2. The DNA sequences used for the preparation of (3+1)-type G-hairpin, G-triplex and G-quadruplexes. Six contiguous Gs were used in this case. The colored regions at the middle are the G-repeat sequences that form the notable structure. The bold letter regions at both the termini indicate the single-stranded regions that are needed to attach these duplexes inside the origami frame. This figure is taken from the supporting information of our recent report [Rajendran *et al.*, *Angew. Chem. Int. Ed.*, **2014**, *53*, 4107.] as we have used the same sequences in this study.

京都大学 KYOTO UNIVERSITY
A Self-archived copy in
Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp
京都大学学術情報リポジトリ KURENAI 紅 Kyoto University Research Information Repository

1) (3+1)-type G-hairpin: Three contiguous Gs



2) (3+1)-type G-triplex: Three contiguous Gs



3) (3+1)-type G-quadruplex: Three contiguous Gs

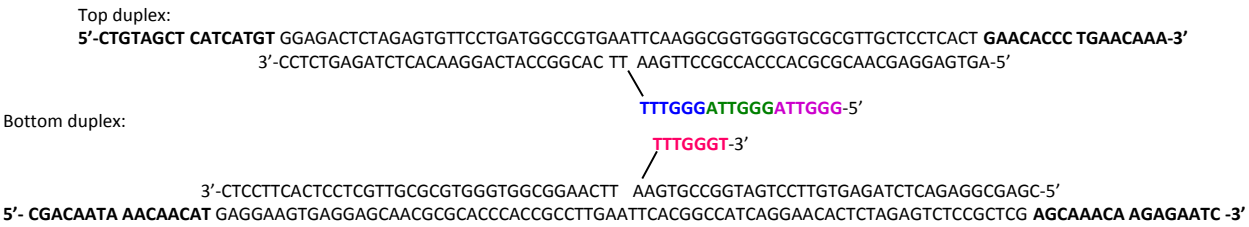


Fig. S3. The DNA sequences used for the preparation of (3+1)-type G-hairpin, G-triplex and G-quadruplexes. Three contiguous Gs were used here. The colored regions at the middle are the G-repeat sequences that form the notable structure. The bold letter regions at both the termini indicate the single-stranded regions that are needed to attach these duplexes inside the origami frame. This figure is taken from the supporting information of our recent report [Rajendran *et al.*, *Angew. Chem. Int. Ed.*, **2014**, *53*, 4107.] as we have used the same sequences in this study.

京都大学 KYOTO UNIVERSITY
A Self-archived copy in
Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp
京都大学学術情報リポジトリ KURENAI 紅 Kyoto University Research Information Repository

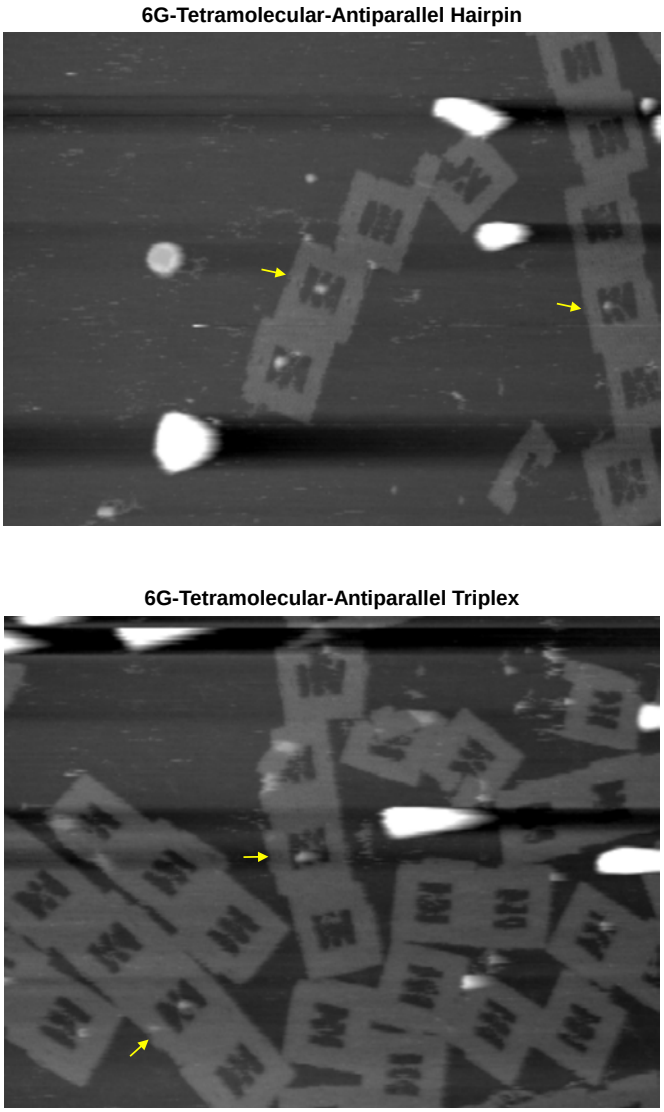


Fig. S4. Representative zoom-out images of the PDC-biotin ligand-induced formation of tetramolecular antiparallel G-hairpin (top) and G-triplex (bottom) structures inside a DNA origami frame. Six contiguous Gs were used in this case. Arrows indicate the X-shaped DNA strands with bound STV inside the origami. Image size: 800 × 600 nm. [Tris-HCl] = 20 mM, pH 7.6; [$MgCl_2$] = 10 mM; [KCl] = 0 mM; [PDC-biotin] = 1 μM; [STV] = 0.2 μM.

京都大学 KYOTO UNIVERSITY
A Self-archived copy in
Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp
京都大学学術情報リポジトリ KURENAI 紅 Kyoto University Research Information Repository

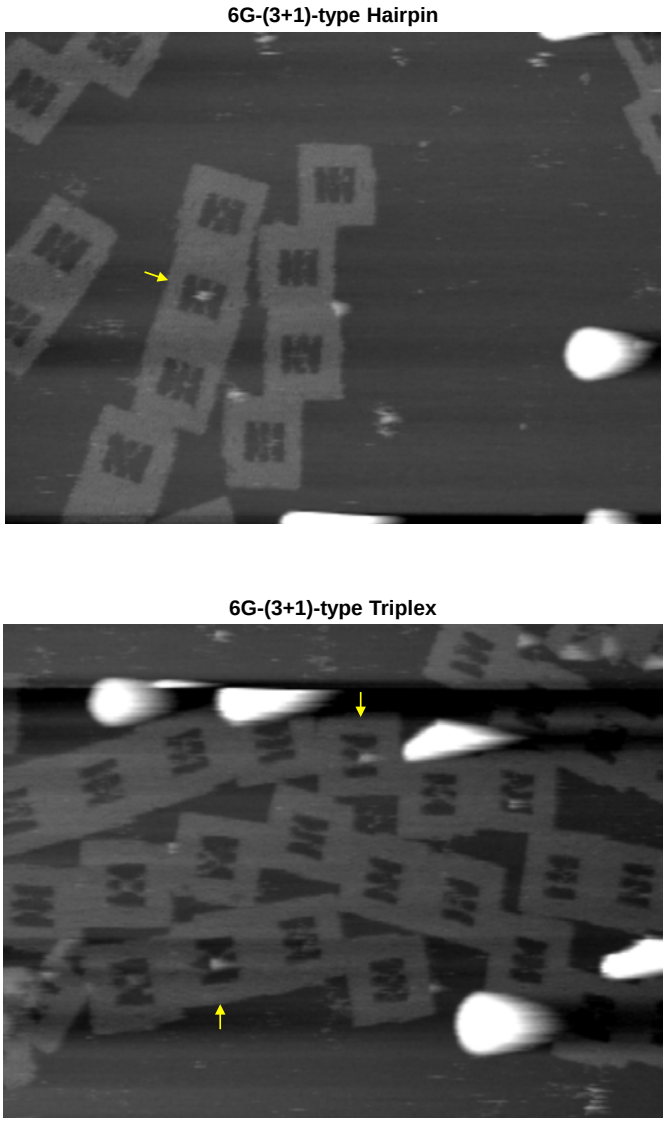


Fig. S5. Representative zoom-out images of the PDC-biotin ligand-induced formation of (3+1)-type G-hairpin (top) and G-triplex (bottom) structures formed inside a DNA origami frame. Six contiguous Gs were used in this case. Arrows indicate the X-shaped DNA strands with bound STV inside the origami. Image size: 800 × 600 nm. [Tris-HCl] = 20 mM, pH 7.6; [$MgCl_2$] = 10 mM; [KCl] = 0 mM; [PDC-biotin] = 1 μM; [STV] = 0.2 μM.

京都大学 KYOTO UNIVERSITY
A Self-archived copy in
Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp
京都大学学術情報リポジトリ KURENAI 紅 Kyoto University Research Information Repository

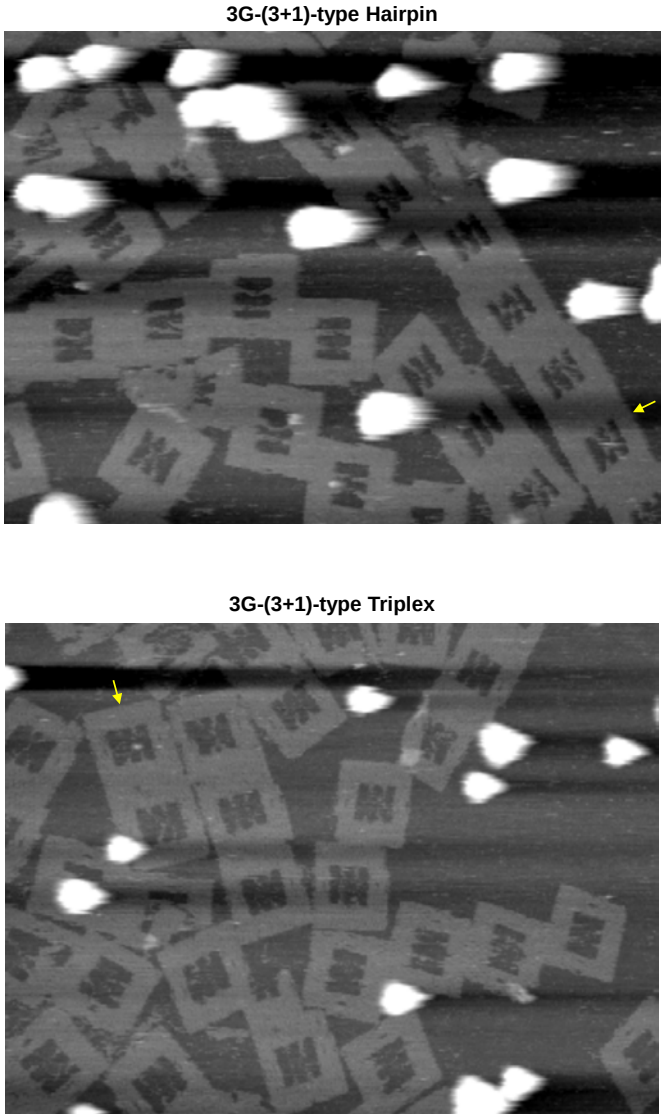


Fig. S6. Representative zoom-out images of the PDC-biotin ligand-induced formation of (3+1)-type G-hairpin (top) and G-triplex (bottom) structures formed inside a DNA origami frame. Three contiguous Gs were used in this case. Arrows indicate the X-shaped DNA strands with bound STV inside the origami. Image size: 800 × 600 nm. [Tris-HCl] = 20 mM, pH 7.6; [$MgCl_2$] = 10 mM; [KCl] = 0 mM; [PDC-biotin] = 1 μM; [STV] = 0.2 μM.